Estimating clinical parameters and risk factors from retinal images using deep learning in the GoDARTS bio-resource

Supervisors:

- Prof. Emanuele Trucco, Prof. Stephen McKenna, Computing (SSEN), University of Dundee
- Dr. Alexander Doney, School of Medicine, University of Dundee
- Dr. V. Prathiba, MDRF, Chennai (India)

Syed Mohammad Ghouse

PhD Student

University of Dundee

Vampire

CVIP DUNDEE

INSPIRED INdia-Scotland Partnership for pRecision mEdicine in Diabetes

GoDARTS GENETICS OF DIABETES AUDIT AND RESEARCH TAYSIDE AND SCOTLAND

Funded by NHS National Institute for Health Research

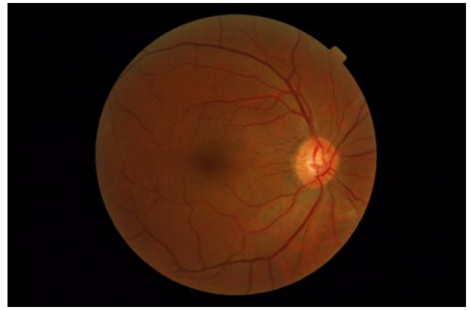
Outline

- Aim of the project
- Dataset
- Image pre-processing
- Convolutional Neural Network
- Results

Aim of Project

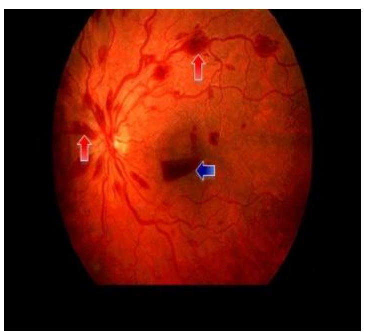
- Investigate the role of the retina as a source of biomarkers for systemic condition in a diabetic cohort (GoDARTS),
  - using deep learning (DL) techniques
  - incorporating explanations of results

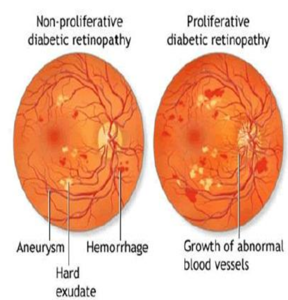
Background

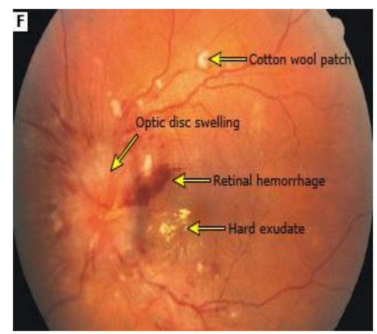


A fundus camera image

- Retina as source of biomarkers for early signs of brain and vascular diseases, e.g. stroke, Alzheimer's disease, diabetes and complications, dementias.







Ref: 1. A. London, I. Benhar, and M. Schwartz. The retina as a window to the brain—from eye research to cns disorders. doi: 10.1038/nrneurol.2012.227.
2. M. Abramoff, M. Garvin, M. Sonka, Retinal imaging and image analysis, IEEE Rev. Biomed. Eng. 3 (2010) 169–208.
3. https://reference.medscape.com/features/slideshow/retina#1
4. http://www.rroij.com/open-access/detection-of-retinal-hemorrhage-in-fundusimages-by-classifying-the-splat-featuresusing-svm.php?aid=50604
5. https://crossfithartford.com/dummies_copper_vs_silver_wiring_eyes.php

University of Dundee
Vampire
CVIP DUNDEE
INSPIRED INdia-Scotland Partnership for pRecision mEdicine in Diabetes
GoDARTS Genetics of Diabetes Audit and Research Tayside and Scotland
Funded by NHS National Institute for Health Research

Challenges

- Retinal biomarkers for systemic diseases (e.g. dementia) may not be directly visible.
- No pre-defined set of rules for patient stratification.

MACE (within 1 year of imaging)

No MACE

Experiments using GoDARTS subset

Dataset

- A total of 1,714 images for the preliminary experiment.
- Features considered for the experiment:
  - Clinical measurements – chol, sbp, dbp, gh, trig, hdl
  - general – age, gender, eversmoked
  - Clinical events – MACE, dementia
- Data is randomly divided as 70% training, 10% validation and 20% testing

University of Dundee
Vampire
CVIP DUNDEE
INSPIRED INdia-Scotland Partnership for pRecision mEdicine in Diabetes
GoDARTS GENETICS OF DIABETES AUDIT AND RESEARCH TAYSIDE AND SCOTLAND
Funded by NHS National Institute for Health Research

Dataset with descriptive statistics

Feature	all_count	Overall mean(std)/prop.	Train_count	Train mean(std)/prop.	Val_count	Val mean(std)/prop.	Test_count	Test mean(std)/prop.
dbp	1713	74.35(7.57)	1199	74.41(7.66)	171	73.92(8.0)	343	74.0(7.28)
sbp	1713	137.82(12.02)	1199	138.13(12.37)	171	137.43(11.0)	343	137.0(11.46)
gh	1713	7.32(1.25)	1198	7.35(1.26)	172	7.26(1.0)	343	7.0(1.26)
chol	1714	4.25(0.77)	1199	4.24(0.76)	172	4.29(1.0)	343	4.0(0.74)
hdl	1714	1.32(0.35)	1199	1.33(0.35)	172	1.28(0.0)	343	1.0(0.33)
trig	1598	2.19(1.33)	1117	2.2(1.29)	162	2.32(1.0)	319	2.0(1.44)
age_at_imaging	1714	70.29(9.59)	1199	70.42(9.61)	172	69.79(9.0)	343	70.0(9.77)
gend(%)	1714	0 56.65 1 43.35	1199	0 56.55 1 43.45	172	0 51.74 1 48.26	343	0 59.48 1 40.52
eversmoker(%)	1714	1 79.7 0 20.3	1199	1 80.15 0 19.85	172	1 81.98 0 18.02	343	1 76.97 0 23.03
MACE(%)	1714	0 57.35 1 42.65	1199	0 57.05 1 42.95	172	0 58.72 1 41.28	343	0 57.73 1 42.27
dementia(%)	1714	0 85.71 21 9.51 22 4.38 20 0.41	1199	0 85.74 21 9.67 22 4.25 20 0.33	172	0 84.30 21 9.88 22 4.65 20 1.16	343	0 86.30 21 8.75 22 4.66 20 0.29
dement_binary(%)	1714	0 85.71 1 14.29	1199	0 85.74 1 14.26	172	0 84.3 1 15.7	343	0 86.3 1 13.7

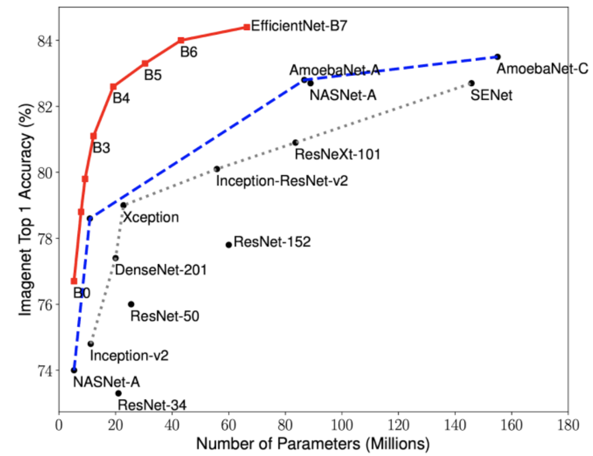
Image Preprocessing

- Images are cropped to obtain the circular region of retina and resized to 512 x 512 (HxW).
- Method 1:
  - Local average color using Gaussian filter is subtracted from the image and mapped to the 50% of gray. [adapted from Kaggle DR winner][1][2]
- Method 2:
  - RGB to YCrCb
  - CLAHE on Y-channel
  - YCrCb to RGB
  - Gamma correction
- For training, validation and testing, images are resized to 260 x 260
- Pixel intensities normalized to 0,1 interval

Ref: 1. https://www.kaggle.com/c/diabetic-retinopathy-detection/discussion/15801
2. Gerrits, N., Elen, B., Craenendonck, T.V. *et al.* Age and sex affect deep learning prediction of cardiometabolic risk factors from retinal images. *Sci Rep* **10,** 9432 (2020). https://doi.org/10.1038/s41598-020-65794-4

Convolutional Neural Network - EfficientNet

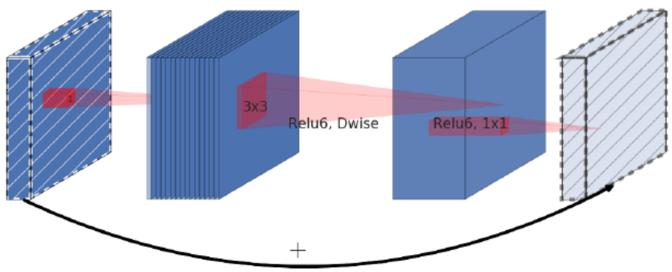
- EfficientNet achieved state of the art performance on Imagenet 2019challenge.
- Main building block:mobile inverted bottleneck, MBConv.
- We modified EfficientNetB2 by adding the GAP + 1FC (sigmoid/linear).
- ~7.7M parameters.



Model size vs Imagenet accuracy

Stage i	Operator $\hat{\mathcal{F}}_i$	Resolution $\hat{H}_i \times \hat{W}_i$	#Channels $\hat{C}_i$	#Layers $\hat{L}_i$
1	Conv3x3	224×224	32	1
2	MBConv1, k3x3	112×112	16	1
3	MBConv6, k3x3	112×112	24	2
4	MBConv6, k5x5	56×56	40	2
5	MBConv6, k3x3	28×28	80	3
6	MBConv6, k5x5	14×14	112	3
7	MBConv6, k5x5	14×14	192	4
8	MBConv6, k3x3	7×7	320	1
9	Conv1x1 & Pooling & FC	7×7	1280	1

EfficientNet-B0 baseline network



Inverted residual block

Ref: EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, http://arxiv.org/abs/1905.11946

University of Dundee

Vampire

CVIP DUNDEE

INSPIRED INdia-Scotland Partnership for pRecision mEdicine in Diabetes

GoDARTS GENETICS OF DIABETES AUDIT AND RESEARCH TAYSIDE AND SCOTLAND

Funded by NHS National Institute for Health Research

CNN training

- Deep learning framework: Keras (TF backend)
- Number of training epochs: 100
- Optimizer: Nadam (lr=0.001)
- Loss: MSE / binary cross-entropy
- Training strategies:
  - Early stopping
  - Model with best validation performance is saved
  - Reduce Learning Rate

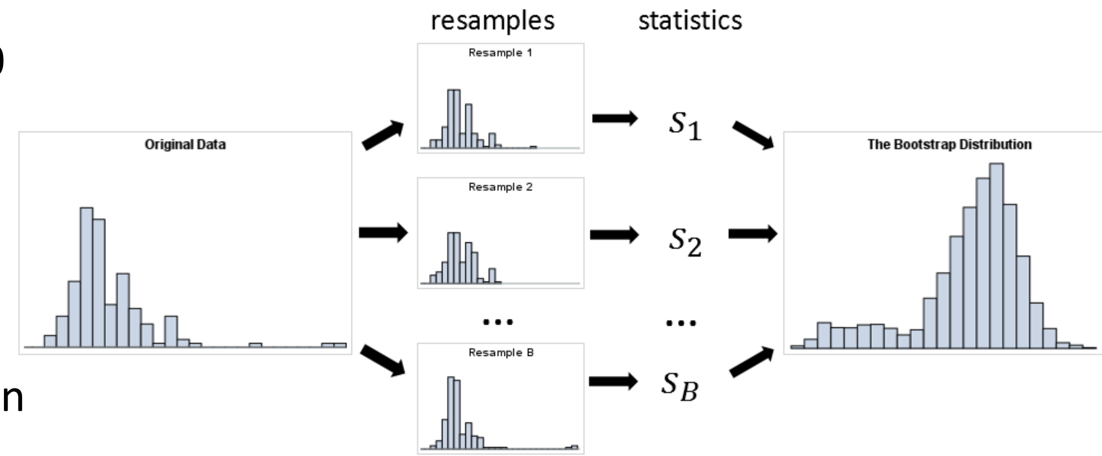
Evaluation on test data - Bootstrap sampling

- Number of bootstrap samples: 2000
- Performance metrics:
  - MAE
  - MAPE
  - R^2
  - AUC
  - Accuracy

For regression

For classification

- 2.5 and 97.5 percentiles as 95% CI are reported



$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| * 100$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)}{\sum_{i=1}^{n}(y_i - \bar{y})}$$

y_i true value of *i*-th sample; $\bar{y}$ mean value; $\hat{y}_i$ predicted value

Figure Ref: https://blogs.sas.com/content/iml/2018/12/12/essential-guide-bootstrapping-sas.html

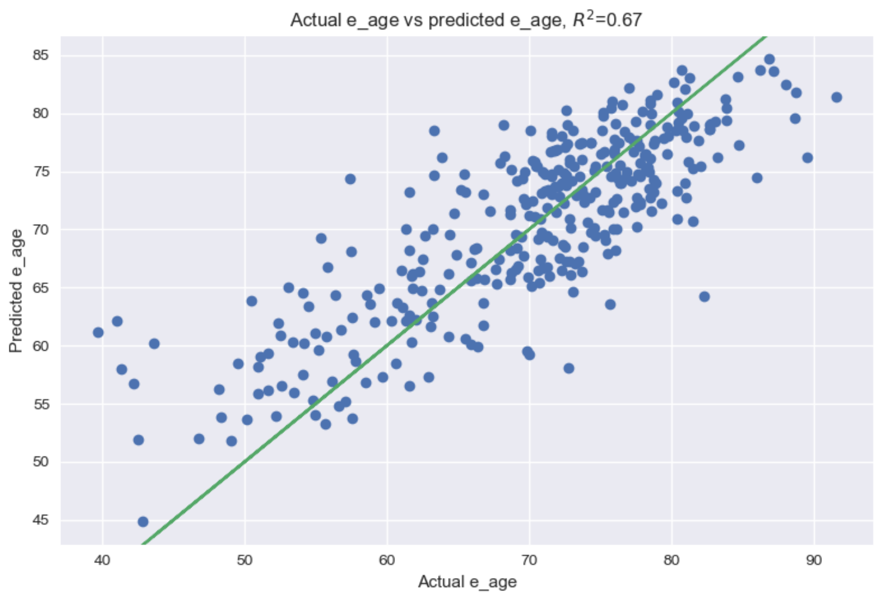
University of Dundee
Vampire
CVIP DUNDEE
INSPIRED India-Scotland Partnership for pRecision mEdicine in Diabetes
GoDARTS Genetics of Diabetes Audit and Research Tayside and Scotland
Funded by NHS National Institute for Health Research

Bootstrap results – continuous features

Feature	Model performance (EfficientNetB2)	Baseline (mean)
Age_at_img: mae(95% CI)	4.37(4.0,4.76)	7.63(6.88,8.42)
Age_at_img: mape(95% CI)	6.7(6.0,7.45)	11.94(10.55,13.43)
Age_at_img: R2(95% CI)	0.66(0.6,0.72)	0
dbp: mae(95% CI)	5.78(5.35,6.25)	5.87(5.43,6.32)
dbp: mape(95% CI)	7.92(7.3,8.57)	8.07(7.43,8.73)
dbp: R2(95% CI)	0.01(-0.09,0.1)	0
sbp: mae(95% CI)	9.14(8.36,9.96)	8.82(8.07,9.62)
sbp: mape(95% CI)	6.9(6.25,7.57)	6.57(5.97,7.21)
sbp: R2(95% CI)	-0.07(-0.17,0.03)	0
gh: mae(95% CI)	0.91(0.83,1.01)	0.92(0.82,1.04)
gh: mape(95% CI)	12.08(11.13,13.05)	12.36(11.07,13.77)
gh: R2(95% CI)	-0.01(-0.06,0.05)	0
chol: mae(95% CI)	0.61(0.56,0.66)	0.59(0.54,0.64)
chol: mape(95% CI)	14.77(13.51,16.07)	14.19(13.0,15.4)
chol: R2(95% CI)	-0.06(-0.13,0.0)	0
hdl: mae(95% CI)	0.25(0.23,0.28)	0.26(0.23,0.28)
hdl: mape(95% CI)	19.98(18.0,21.9)	20.5(18.54,22.53)
hdl: R2(95% CI)	0.0(-0.08,0.08)	0
trig: mae(95% CI)	0.97(0.87,1.1)	0.9(0.76,1.06)
trig: mape(95% CI)	59.62(53.64,66.15)	53.25(46.9,60.35)
trig: R2(95% CI)	-0.04(-0.15,0.02)	0

Model performance on age group - test data

Age group	count	mean_actual_age	mean_predict_age	MAE	MAPE
0-10	0	-	-	-	-
10-20	0	-	-	-	-
20-30	0	-	-	-	-
30-40	1	39.663	61.137	21.474	54.141
40-50	11	45.024	55.129	10.105	23.171
50-60	43	55.080	60.109	5.719	10.490
60-70	77	65.432	67.493	4.112	6.317
70-80	168	74.574	73.737	3.506	4.724
80-90	42	82.925	78.440	4.912	5.886
90-100	1	91.606	81.432	10.174	11.106



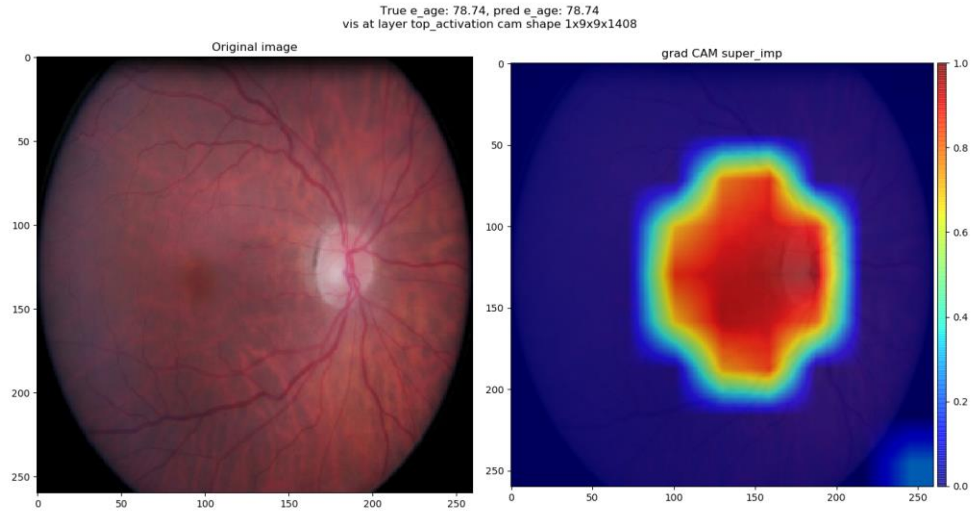
University of Dundee

GoDARTS GENETICS OF DIABETES AUDIT AND RESEARCH TAYSIDE AND SCOTLAND

Funded by NHS National Institute for Health Research

Visualization: age estimation

- <u>Grad-CAM</u> algorithm
- Visualization results at the top activation layer
- Heatmap at activation layer resized from 9x9 to 260 x 260.
- <u>Optic disc and macula are selected as important for age estimation</u>.
- Further investigation needed.



University of Dundee

Vampire

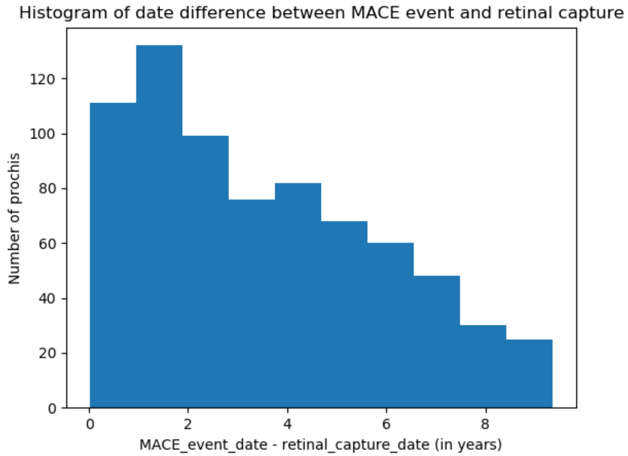
Funded by NHS National Institute for Health Research

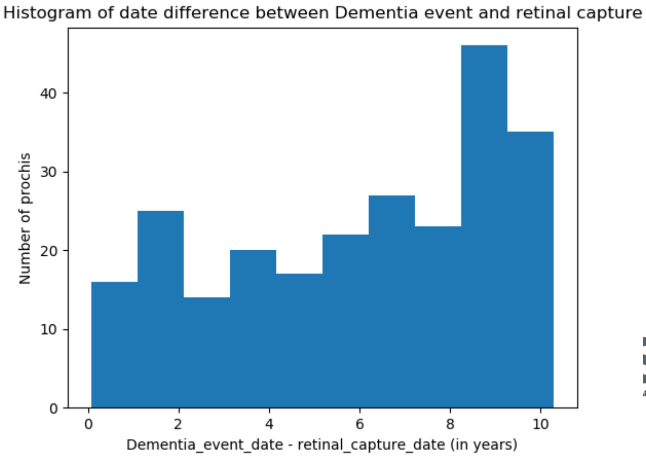
Bootstrap results – categorical features

Feature	Model performance (EfficientNetB2)
gender: AUC(95% CI)	0.75(0.7,0.8)
gender: Accuracy(95% CI)	0.62(0.57,0.67)
eversmoked: AUC(95% CI)	0.56(0.49,0.63)
eversmoked: Accuracy(95% CI)	0.77(0.73,0.81)
MACE: AUC(95% CI)	0.44(0.38,0.51)
MACE: Accuracy(95% CI)	0.53(0.48,0.58)
dementia_binary: AUC(95% CI)	0.62(0.53,0.7)
dementia_binary: Accuracy(95% CI)	0.85(0.81,0.88)

Note: accuracies are high due to class imbalance in the test dataset

- Histogram of time lapse (dates difference) btw. retinal image capture and event





Funded by NHS National Institute for Health Research

Related work

- Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning (nature biomedical engineering, Feb. 2018)[1]
  - Dataset - UK biobank and EyePACS (1.8M images from 284k individuals)
  - Mean age is 55 years in training

- Effects of hypertension, diabetes, and smoking on age and sex prediction from retinal fundus images (Scientific Reports, March 2020)[2]
  - Dataset – SBRIA (412k images from 155K)
  - Mean age is 47 years in training data

- Age and sex affect deep learning prediction of cardiometabolic risk factors from retinal images (scientific reports(nature research), July 2020)[3]
  - Dataset – Qatar Biobank subset (12k images from 3k individuals)
  - Mean age is 40 years on whole dataset

Ref: 1. https://doi.org/10.1038/s41551-018-0195-0
2. https://doi.org/10.1038/s41598-020-61519-9
3. https://doi.org/10.1038/s41598-020-65794-4

Experiment: I2x event classification in GoDARTS

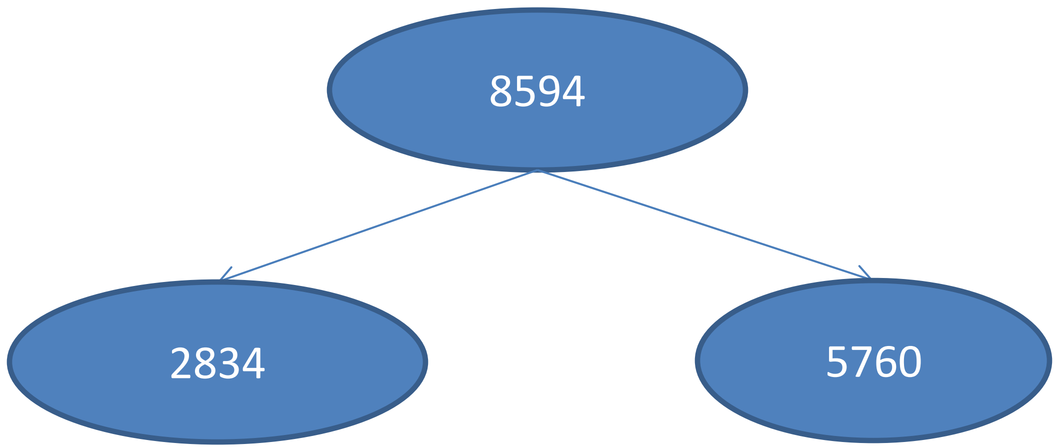
Selection of events

- For this experiment, selected ICD-10 codes:
  - I20-I25 Ischaemic heart diseases
  - I26-I28 pulmonary heart disease and diseases of pulmonary circulation
- Individuals admitted to hospital with the above ICD-10 codes are filtered from the database
- Question: Whether retina image can stratify I2x event form non-I2x event?

Ref: https://icd.who.int/browse10/2015/en#

Data collection

- In total 102,455 retinal images are available in GoDARTS from 8,594 individuals.



- Admitted to hospital due to I2x event
- 369 died in hospital
- Filtered based on the last discharge date

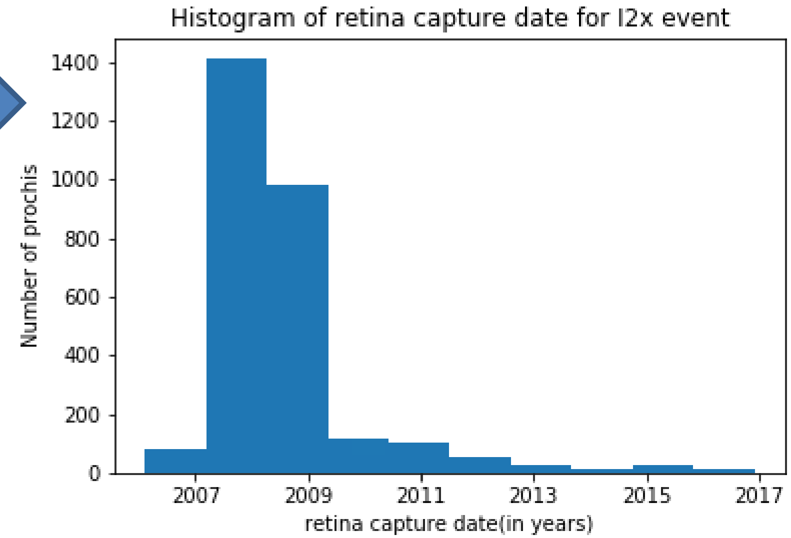
- Not found any hospital admission due to I2x event

Retinal image selection

- For I2x event:
  - The most recent retinal image available when the individual is last admitted to hospital.
- For non-I2x event:
  - Case 1: The most recent retinal image available.
  - Case 2: The very first retinal image available.
- Two datasets prepared:
  - (I2x event, case1) - Recent
  - (I2x event, case2) - First

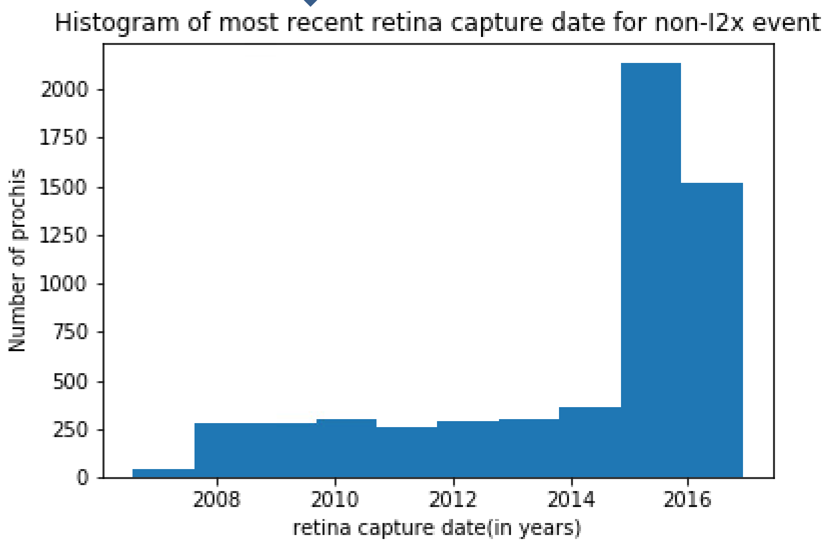
Histograms of retina capture dates

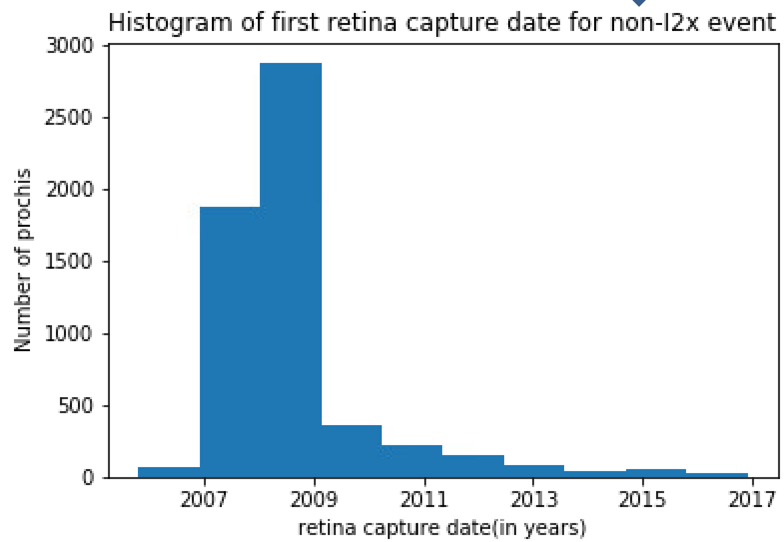
Individuals with I2x event



Individuals with non-I2x event – Case 1

Individuals with non-I2x event – Case 2





Results

- Randomly split the data into 70% training, 10% validation and 20% testing.

Category	I2x (count)	Non-I2x (count)	Total count
Train	2003 (33.3%)	4012 (66.7%)	6015 (100%)
Val	301(35%)	559 (65%)	860 (100%)
Test	530 (30.83%)	1189 (69.17%)	1719 (100%)

- Training - EfficientNetB2 for 10 epochs.
- Results on test data:

Category	AUC	Accuracy
Case 1 (Recent)	91.46%	84.8%
Case 2 (First)	62.05%	68.6%

- Thorough Investigation is required.

University of Dundee Vampire CVIP DUNDEE INSPIRED INdia-Scotland Partnership for pRecision mEdicine in Diabetes GoDARTS GENETICS OF DIABETES AUDIT AND RESEARCH TAYSIDE AND SCOTLAND Funded by NHS National Institute for Health Research

Future work

- Replicate the experiments over a large volume of GoDARTS dataset.
- Thorough investigation on the interpretability of CNN.
- Focus on a narrow set of clinical conditions.

University of Dundee

Vampire

CVIP DUNDEE

INSPIRED
INdia-Scotland Partnership for pRecision mEdicine in Diabetes

GODARTS
GENETICS OF DIABETES AUDIT AND RESEARCH TAYSIDE AND SCOTLAND

Funded by NHS National Institute for Health Research

Acknowledgements

- Dr. Muthu Mookiah, Stephen Hogg – from computing, UoD.
- INSPIRED team especially, Aravind and Anand.
- HIC team especially, Dr. Joseph Ward.

Funders:

- The research was commissioned by the National Institute for Health Research using Official Development Assistance (ODA) funding [INSPIRED 16/136/102].

- Disclaimer : The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

University of Dundee

Vampire

CVIP DUNDEE

INSPIRED INdia-Scotland Partnership for pRecision mEdicine in Diabetes

GoDARTS GENETICS OF DIABETES AUDIT AND RESEARCH TAYSIDE AND SCOTLAND

Funded by NHS National Institute for Health Research

Thank You!